By J. Michael McWilliams, Gabe Weinreb, Lin Ding, Chima D. Ndumele, and Jacob Wallace

DOI: 10.1377/hlthaff.2022.00916
HEALTH AFFAIRS 42,
NO. 1 (2023): 105–114

Risk Adjustment And Promoting Health Equity In Population-Based Payment: Concepts And Evidence

ABSTRACT The objective of risk adjustment is not to predict spending accurately but to support the social goals of a payment system, which include equity. Setting population-based payments at accurate predictions risks entrenching spending levels that are insufficient to mitigate the impact of social determinants on health care use and effectiveness. Instead, to advance equity, payments must be set above current levels of spending for historically disadvantaged groups. In analyses intended to guide such reallocations, we found that current risk adjustment for the community-dwelling Medicare population overpredicts annual spending for Black and Hispanic beneficiaries by $376–$1,264. The risk-adjusted spending for these populations is lower than spending for White beneficiaries despite the former populations' worse risk-adjusted health and functional status. Thus, continued movement from fee-for-service to population-based payment models that omit race and ethnicity from risk adjustment (as current models do) should result in sizable resource reallocations and incentives that support efforts to address racial and ethnic disparities in care. We found smaller overpredictions for less-educated beneficiaries and communities with higher proportions of residents who are Black, Hispanic, or less educated, suggesting that additional payment adjustments that depart from predictive accuracy are needed to support health equity. These findings also suggest that adding social risk factors as predictors to spending models used for risk adjustment may be counterproductive or accomplish little.

J. Michael McWilliams (mcwilliams@hcp.med.harvard.edu), Harvard University and Brigham and Women's Hospital, Boston, Massachusetts.

Gabe Weinreb, Harvard University.

Lin Ding, Harvard University.

Chima D. Ndumele, Yale University, New Haven, Connecticut.

Jacob Wallace, Yale University.

Population-based payment systems distribute spending on the basis of population characteristics, unlike fee-for-service systems, in which spending is distributed according to service use. Accordingly, population-based payment models, as in the Medicare Shared Savings Program or Medicare Advantage (MA) program, can facilitate the resource reallocations necessary to address health care disparities. Risk adjustment is the mechanism by which payment is allocated in these models.

Traditionally, risk adjustment has been conceived and executed purely as a predictive exercise. Regression is used to predict total annual per person spending as a function of demographic and clinical characteristics. A person's predicted spending is converted to a risk score, which is applied to a base regional rate to determine the prospective payment or benchmark for that person. The more accurately spending is predicted (that is, the better the fit of the regression model), the more closely payment matches spending, thereby equalizing financial risk across providers or plans serving different populations and limiting incentives to attract favorable risks (patients with overpredicted spending) or avoid unfavorable risks (patients with underpredicted spending).

A commonly voiced concern with the transition to population-based payment is that risk adjustment will fail to account for historically marginalized groups' presumed higher spending, thereby exacerbating health disparities. Framed as solving a prediction problem, social risk adjustment is thus often thought to achieve its goal by adding social risk factors as predictors to standard risk-adjustment models. Various studies have considered the incremental predictiveness of measurable social risk factors and made recommendations about which to add, but this line of research has focused largely on health outcomes or acute care use, as opposed to spending.[1] Implicit in many calls for "improved" risk adjustment is an assumption that social risk factors predict higher spending, that the problem is their omission from predictive models, and that equity-promoting reallocations thus can be motivated by predictive accuracy.

However, attempting to support more equitable care by improving the predictive accuracy of risk adjustment is a fundamentally limited strategy because historical and current levels of spending (the target of prediction) are unlikely to be the desired levels of spending for those populations. People who experience social disadvantage may use less health care and have lower spending than others with the same clinical needs.[2–4] For example, they may have less income to spend on health care, have less generous insurance coverage, be less aware of their health care needs because of lower educational attainment, face greater barriers to obtaining care (for example, travel and time constraints), or encounter additional barriers from other manifestations of structural or interpersonal racism. The inclusion of markers of social disadvantage in risk-adjustment models may therefore improve predictive accuracy but reduce payments for underserved populations relative to models that omit these markers.

Moreover, current spending for historically marginalized groups may be too low to support equitable care because providers serving those groups may have insufficient resources to improve the quality of care or provide the additional supportive services (for example, case management) necessary to mitigate the adverse impact of social determinants on health care use and effectiveness. Many supportive services are not reflected in fee-for-service spending.

Thus, even if the addition of some social factors to standard risk adjustment results in higher population-based payments for populations with a higher prevalence of those factors, the adjustments merely recover spending levels under fee-for-service that are believed to be too low to cover the costs of reducing disparities. To address disparities, payment must instead be set above current spending (or an accurate prediction thereof).

More generally, the objective of risk adjustment is not solely to predict observed spending accurately. Rather, it is to support the broader social goals of payment reform—to make the health care system more efficient and equitable.[5,6] With that as the goal, current spending is inherently the wrong target for population-based payments. A reformed system should encourage the desired level and distribution of spending, not entrench the status quo.

That risk adjustment presents trade-offs between fit (the predictive accuracy of a model) and other objectives has been well described. Improving fit inherently weakens the power of incentives in a population-based payment system.[5] As payments (or benchmarks) are adjusted for more markers of health care use (for example, diagnoses) or for use directly (for example, lagged indicators of hospitalization), risk-bearing entities save less from curbing unnecessary or avoidable care (reducing use reduces payment). In the extreme, adjusting for use of each service would achieve perfect fit but revert payment incentives to those of fee-for-service. To some extent, incentives encouraging risk selection (deficient fit) must be tolerated to allow the payment system to control spending.

Likewise, setting population-based payments (or benchmarks) above an accurate prediction of fee-for-service spending for historically disadvantaged groups worsens fit but advances the goal of health equity by mitigating resource disparities that contribute to health disparities and better aligning payment with health care needs (including unmet needs). Deliberately paying above current spending for those groups also protects socially vulnerable patients with underpredicted clinical needs against risk selection and creates incentives for competing providers or plans to attract the underserved with enhanced benefits or services. Several approaches have been developed to set population-based payments at desired rather than accurately predicted levels.[6] Yet concerns about inadequate accounting for social determinants in population-based payment models remain largely framed around the predictive accuracy of standard risk-adjustment methods, often appealing to the promise of advanced prediction tools, such as machine learning and artificial intelligence, in proposed solutions.[7–10]

To inform payment policy intended to support health equity, in this study we first added individual-level predictors of social disadvantage (race, ethnicity, and educational attainment) to the Hierarchical Condition Categories (HCC) model currently used to risk-adjust payments in Medicare Advantage and benchmarks in the Medicare Shared Savings Program. The results describe existing underpredictions or overpredictions by the HCC model of spending by race, ethnicity, and education. Second, we calculated the associated reallocations across groups achieved by moving from fee-for-service to a population-based payment system under current risk-adjusted methods, which omit these social characteristics as predictors. These reallocations equivalently describe the incentives for a risk-bearing entity to attract people with these characteristics. Third, we compared the HCC-adjusted differences in spending between groups with HCC-adjusted differences in self-reported health status, functional limitations, and access to care to gauge the extent to which reallocations under current risk adjustment are commensurate with addressing evident disparities. Fourth, we compared results when using area-level, instead of individual-level, versions of the same predictors. Finally, we considered the implications of our findings for the targeting and implementation of population-based payment adjustments that depart from predictive accuracy to support health equity.

Study Data And Methods

STUDY POPULATION AND DATA We analyzed Medicare claims from the period 2012–17 for 20 percent annual random samples of fee-for-service beneficiaries and for respondents to the 2012–17 fee-for-service Medicare Consumer Assessment of Healthcare Providers and Systems (CAHPS) surveys. Fee-for-service Medicare CAHPS is administered annually to a national cross-sectional sample of fee-for-service beneficiaries. It assesses patients' experiences with care and collects sociodemographic and health information not available in Medicare administrative data files. The average CAHPS response rate among beneficiaries meeting inclusion criteria was 42.2 percent during our study period. We limited each annual 20 percent or CAHPS sample to beneficiaries continuously enrolled in Parts A and B of fee-for-service Medicare in both the study year (while alive for decedents) and the preceding year (to collect diagnoses for the prospective HCC model). For consistency across samples, we excluded long-term nursing facility residents and beneficiaries with end-stage renal disease from our main analyses, as their numbers are small among the largely community-dwelling CAHPS respondents. We conducted separate analyses of these groups for outcomes available for the 20 percent sample.

STUDY VARIABLES We calculated total annual Medicare spending per beneficiary by summing payments across all services reimbursed by Part A or B. From survey data for CAHPS respondents, we assessed indicators of compromised health or access to care: fair or poor general health status; fair or poor mental health; difficulty with one or more activities of daily living (ADLs); and difficulty accessing routine, urgent, or specialty care in a timely fashion, as defined by a report of never or sometimes receiving care as soon as needed (versus usually or always).

For both the 20 percent and CAHPS samples, we assessed beneficiaries' race and ethnicity from the Research Triangle Institute (RTI) race and ethnicity variable in the Medicare Master Beneficiary Summary File.[11] We focused our main analyses on Black, Hispanic, and non-Hispanic White beneficiaries (categories as defined by the RTI variable); the RTI variable exhibits stronger concordance with self-reported classification for these groups than others.[12] For CAHPS respondents, we additionally assessed educational attainment (dichotomized as less than a high school diploma versus a high school diploma or more) and self-reported race and ethnicity to explore the validity of estimates derived from the RTI variable.

Using 2015–19 data from the Census Bureau's American Community Survey, we created analogues of these variables at the census block group level (proportion of residents without a high school diploma, proportion Black, and proportion Hispanic). Finally, from the Master Beneficiary Summary File, we determined beneficiaries' age, sex, dual eligibility for Medicaid, original reason for Medicare eligibility (age versus qualifying disability), county of residence, and nine-digit ZIP code of residence for linking block group–level variables.

STATISTICAL ANALYSIS We fit a linear regression model of total annual per beneficiary Medicare spending as a function of age, sex, HCC indicators, enrollment segment (aged non-dual-eligible beneficiaries, aged dual-eligible beneficiaries, disabled non-dual-eligible beneficiaries, and disabled dual-eligible beneficiaries), interactions between segments and the HCC indicators, and the added predictor of interest (race and ethnicity or education). Whereas typically the HCC model (a linear regression model) is fit within each segment, this pooled model provided average estimates across community-dwelling segments (see the Supplementary Methods in the online appendix for a discussion of model choice).[13]

We also included county fixed effects in the model because MA payments and the base rates for the regional component of accountable care organization (ACO) benchmarks are set at the county level as a function of past fee-for-service spending in the county. HCC risk scores are applied to county base rates to determine payments or benchmarks. Without adjustment for county, we might erroneously have concluded, for example, that HCC-adjusted benchmarks would undercompensate ACOs for a group disproportionately living in high-spending counties.

Thus, the model estimated within-county differences in HCC-adjusted fee-for-service spending between groups with different race, ethnicity, or educational attainment. From these differences we calculated the extent to which the HCC model, which omits these characteristics, over- or underpredicts fee-for-service spending for each group. These over- or underpredictions characterize the selection incentives and payment allocations that an ACO operating in a given county would face when serving different groups of beneficiaries, assuming that the ACO's benchmark is based entirely on an HCC-adjusted regional rate, as in Medicare Advantage. (At this time, an ACO's benchmark is based on a blend of a risk-adjusted regional component and the ACO's historical spending.) In turn, these estimates show how moving from fee-for-service to a risk-adjusted population-based payment model would reallocate resources across MA plans or ACOs serving different groups. If the HCC model overpredicts spending for historically disadvantaged groups, moving toward risk-adjusted population-based payment would increase payment for them (and vice versa if the model underpredicts spending). Our approach assumed that coefficients for HCCs were similar with and without county fixed effects (appendix exhibit 1).[13]

We then fit the same model to each health status and access-to-care indicator. These models aided normative interpretation of the results for spending. If, for example, beneficiaries with low education have lower HCC-adjusted spending but worse HCC-adjusted health status or access to care, we can surmise that their lower spending is not commensurate with their health and needs and that higher spending could help address disparities. All analyses used robust variance estimators, clustered at the county level. Analyses of CAHPS data additionally applied survey weights to account for nonresponse.

LIMITATIONS Our study had several limitations. First, we relied on fee-for-service Medicare claims. Although our estimates are informative for understanding how the HCC model allocates payment and creates selection incentives across groups of MA enrollees, the estimates would differ somewhat if they were based on MA data. For example, spending for beneficiaries with less education may be lower in traditional fee-for-service Medicare than in Medicare Advantage, in part because beneficiaries with less education may be less likely to have supplemental insurance in traditional Medicare. We might therefore expect smaller spending differences in Medicare Advantage. In a sensitivity analysis of the fee-for-service CAHPS sample, we additionally adjusted for self-reported sources of supplemental coverage to better approximate differences in spending between groups that we would observe in Medicare Advantage.

Adding social factors, particularly race and ethnicity, to the HCC model can entrench health disparities instead of reducing them.

Second, although our analysis showed the extent to which the current (HCC) risk-adjustment system over- or underpredicts fee-for-service spending for historically marginalized groups and communities, it could not determine the socially optimal payment level. That depends on social values, the extent of underspending for the underserved, and the extent to which payment increases would be passed through by providers or plans to populations in need. Nevertheless, such uncertainty should not hinder the initiation of a desirable direction for payment, and our estimates inform where adjustments are needed to increase payment above predictions made by the HCC model. For example, if the model predicted spending accurately for a group reporting worse access and health, we would consider a payment increase above the predicted level for that group.

Third, the social characteristics we examined were limited to race, ethnicity, and education, selected because they could be ascertained at both individual and block group levels. These are powerful predictors of disadvantage mediated by a range of mechanisms and strongly correlated with other markers; conclusions were similar, for example, in analyses using the Area

Deprivation Index. Moreover, our study was a proof-of-concept analysis that produced an instructive set of varied results and implications across the groups studied. It was not intended to be comprehensive in the predictors examined, as the objective was not to predict better but, rather, to illustrate conceptual and empirical considerations underlying sound payment policy.

Finally, although our analysis can inform payment reallocations to support health equity, it did not assess the extent to which reallocated resources reached specific populations, as intended, to improve their care.

Study Results

STUDY POPULATION The sociodemographic characteristics of the 2012–17 CAHPS sample were similar to those of the 20 percent sample of community-dwelling Medicare beneficiaries (exhibit 1).

SPENDING After adjustment for age, sex, enrollment segment, HCCs, and county, total annual Medicare spending per beneficiary was $574 lower for Black beneficiaries and $1,462 lower for Hispanic beneficiaries than for White beneficiaries in the 20 percent samples (exhibit 2). These estimates suggest substantial overprediction of fee-for-service spending for Black and Hispanic beneficiaries by the current or

EXHIBIT 1

Characteristics of the fee-for-service Medicare Consumer Assessment of Healthcare Providers and Systems (CAHPS) and Medicare 20 percent samples, 2012–17

Beneficiary characteristics	CAHPS sample (*N* = 512,401)	Medicare 20% sample (*N* = 32,721,400)
Mean age, years	72.9	72.5
Sex, %		
Female	55.0	55.0
RTI race and ethnicity variables, %		
American Indian or Alaska Native	0.5	0.5
Asian or Pacific Islander	2.3	2.3
Black	8.5	8.9
Hispanic	4.9	5.4
Non-Hispanic White	82.2	81.4
CAHPS self-reported race and ethnicity, %		
American Indian or Alaska Native	1.8	—[a]
Asian or Pacific Islander	2.7	—[a]
Black or African-American	7.7	—[a]
Hispanic or Latino	5.1	—[a]
Multiracial	0.4	—[a]
White	78.6	—[a]
Education, %		
No high school diploma	13.3	—[a]
Block group measures, mean %		
Percent without a high school diploma	13.9	14.0
Percent Black	10.1	10.4
Percent Hispanic	12.1	12.4
Enrollment segments, %		
Aged, dual eligible	6.2	7.1
Aged, non–dual eligible	72.3	68.6
Disabled, dual eligible	9.6	11.6
Disabled, non–dual eligible	12.0	12.7
Total annual Medicare spending per beneficiary, $	8,506	9,000

SOURCE Authors' analysis of enrollment data from the Medicare Master Beneficiary Summary File for the period 2012–17 for fee-for-service CAHPS survey respondents and 20 percent samples of fee-for-service beneficiaries, and of fee-for-service CAHPS survey data. **NOTES** Sample sizes are measured in beneficiary-years. Descriptive statistics for CAHPS variables are weighted using CAHPS survey weights. The Research Triangle Institute (RTI) race and ethnicity variable is an enhanced version of the base Master Beneficiary Summary File race and ethnicity data that uses surname and geographic analysis to improve accuracy for Hispanic and Asian populations; classification of Black beneficiaries by both variables is based on self-reported data collected by the Social Security Administration. To support consistent comparisons with the RTI variable classification, all CAHPS respondents who self-identified as Hispanic or Latino in response to an item about Hispanic or Latino descent were categorized as Hispanic or Latino. Those who self-identified as White and one other category of race were assigned to the non-White category. Those reporting two different non-White categories were classified as multiracial. Estimates from our main analyses for Black and Hispanic beneficiaries were not appreciably changed by alternative categorizations of the CAHPS responses in sensitivity analyses. [a]Not applicable.

EXHIBIT 2

Medicare spending differences, by beneficiary and community-level characteristics, after standard risk adjustment, 2012–17

Characteristics	HCC-adjusted difference in per beneficiary annual Medicare spending, $	95% CI
RTI race and ethnicity		
Non-Hispanic White	Ref	
Black	−574	−691, −456
Hispanic	−1,462	−1,609, −1,315
Education		
High school diploma or more	Ref	
No high school diploma	−85	−289, 120
Block group measures[a]		
Percent without a high school diploma	−45	−68, −23
Percent Black	−7	−49, 36
Percent Hispanic	−115	−151, −79

SOURCE Authors' analysis of fee-for-service Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey data, Medicare claims and enrollment data, and American Community Survey data. **NOTES** Estimates for individual-level race and ethnicity and block group–level characteristics are from analysis of the Medicare fee-for-service 20 percent samples, whereas estimates for individual-level education are from analysis of the CAHPS sample. The Research Triangle Institute (RTI) race and ethnicity designation method is described in the exhibit 1 notes. HCC is Hierarchical Condition Categories. [a]Estimates for each block group measure have been rescaled to reflect a 1-standard-deviation change in the measure. For example, an increase of 1 standard deviation in the block group proportion of residents without a high school diploma is associated with a $45 decrease in per beneficiary Medicare spending.

standard (HCC) model, which does not include race or ethnicity. In turn, population-based payments set by applying HCC risk scores to a county base rate would redistribute payment away from White beneficiaries (−$198 per beneficiary) toward Black (+$376) and Hispanic (+$1,264) beneficiaries (exhibit 3). These payment reallocations equivalently quantify the relative selection incentives that an ACO receiving such risk-adjusted population-based payments would face, on average, in a given county; the ACO would have a strong incentive to attract Hispanic and Black residents of that county. Conversely, adding race and ethnicity to the HCC model would lower payments for Black and Hispanic beneficiaries (but would improve the predictive accuracy [fit] of the model).

Estimates were less precise but were substantively similar in analyses of the CAHPS sample when we used the RTI variable (appendix exhibit 2),[13] suggesting that analyses of other variables available only for CAHPS respondents also should generalize to the full community-dwelling fee-for-service Medicare population. Within the CAHPS sample, estimates also were substantively similar when we assessed race and ethnicity using the RTI variable versus self-reports (appendix exhibit 2).[13] This supplementary analysis additionally revealed that the HCC model substantially overpredicts fee-for-service spending for Asian or Pacific Islander beneficiaries. Estimates for American Indian or Alaska Native beneficiaries varied across samples and data sources, limiting inferences.

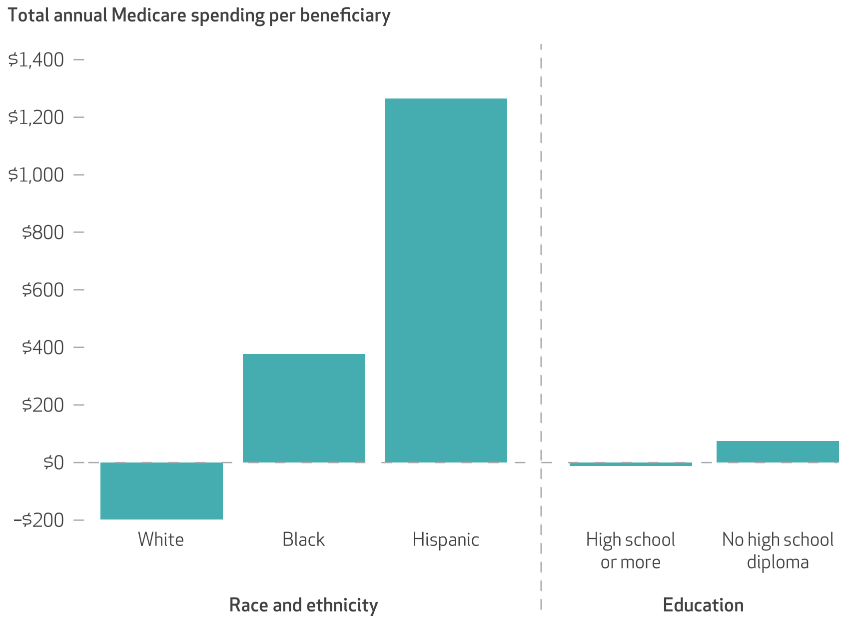
In contrast to findings for race and ethnicity, HCC-adjusted spending was minimally lower for beneficiaries with less than a high school diploma than for those with more education in the same county (−$85 per beneficiary; 95% confidence interval: −289, 120) (exhibit 2), implying only a small equity-promoting reallocation (+$73) from moving away from fee-for-service toward population-based payment under the HCC model (exhibit 3). Similarly, census block group aggregates of race, ethnicity, and education did not predict spending that differed markedly from what the HCC model (which, again, does not include these variables) would predict when applied to a county base rate. As detailed in exhibit 2, an increase in the proportion of residents without a high school diploma equal to a full standard deviation in the block group distribution predicted $45 lower spending per beneficiary. Similarly scaled estimates for the proportion of residents who were Black and Hispanic were −$7 and −$115. These findings for area-level predictors suggest that moving from fee-for-service to population-based payment under current risk adjustment would result in minimal to modest reallocations toward communities with higher proportions of Black, Hispanic, or less-educated residents and would thus give ACOs (or, by extension, MA plans) minimally to modestly stronger incentives to enter or expand their provider networks in those communities relative to other communities.

HEALTH STATUS, FUNCTIONAL STATUS, AND ACCESS TO CARE Despite lower HCC-adjusted spending, Black and Hispanic beneficiaries reported significantly worse general and mental health status and more difficulties with ADLs than White beneficiaries in HCC-adjusted comparisons (exhibit 4). Disparities in health and functional status were even greater between beneficiaries with versus without a high school diploma, despite smaller spending differences. For example, beneficiaries without a high school diploma were 10.4 percentage points more likely to report being in fair or poor health (sample mean, 29.3 percent) and 4.3 percentage points more likely to report difficulty with an ADL (mean, 37.5 percent). Black and Hispanic beneficiaries and those with less than a high school diploma also all reported worse access to care (exhibit 4). Findings were mostly similar for Asian or Pacific Islander and American Indian or Alaska Native beneficiaries (appendix exhibit 3).[13] Associations between block group variables and health or functional status were mostly similar in direction but were smaller in magnitude. Estimates from logistic regression models were similar (appendix exhibit 4).[13]

SUPPLEMENTARY ANALYSES Spending estimates by race and ethnicity at the individual level were directionally similar across each community-dwelling population segment when analyzed separately, but they differed for Black beneficiaries in the end-stage renal disease population and for both Black and Hispanic beneficiaries in the long-term nursing facility resident population (appendix exhibit 5).[13] Appendix exhibit 6 summarizes three findings.[13] First, estimates were substantively similar after we removed HCCs from the spending model (that is, adjusting only for age, sex, enrollment segment, and county). Second, removing county fixed effects (not included in the HCC model) confirmed that adding indicators of historically disadvantaged groups to the model would generally result in lower or minimally higher payments for them. Third, supplementary analyses using the Area Deprivation Index supported the conclusions from our main analyses. Results of CAHPS analyses were qualitatively similar after adjustment for supplemental insurance (appendix exhibit 7).[13]

EXHIBIT 3

Implied per beneficiary payment redistribution resulting from a transition to population-based payments under standard risk adjustment in Medicare, 2012–17



SOURCE Authors' analysis of fee-for-service Medicare claims and enrollment data and fee-for-service Medicare Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey data. **NOTES** Estimates for individual-level race and ethnicity are from analysis of the 20 percent fee-for-service Medicare samples, whereas estimates for individual-level education are from analysis of the CAHPS sample. Redistributions reflect the extent to which the Hierarchical Condition Categories (HCC) model over- or underpredicts spending for each group and were calculated as follows. The estimates in exhibit 2 describe the HCC-adjusted difference in spending between groups. To describe the difference between an HCC-adjusted population-based payment (the average risk-adjusted spending for all groups) and spending for a single group, we applied the group population shares from exhibit 1 to the estimate in exhibit 2. For example, beneficiaries without a high school diploma have risk-adjusted spending that is $85 lower than those with more education (exhibit 2). Average risk-adjusted spending is a weighted average of spending for those with more and less education according to the distribution in exhibit 1 (86.7% have a high school diploma or more). Thus, the average is 0.867 × $85 = $73.7 higher (the estimate in exhibit 3 within rounding error).

Discussion

In this study of community-dwelling fee-for-service Medicare beneficiaries, Medicare spending was similar or substantially lower for groups at higher risk of experiencing social disadvantage after adjustment for variables in the current HCC risk-adjustment model. That HCC-adjusted spending was not higher for these groups is consistent with the findings of other studies, but it may run counter to expectations.[3,4] For example, some may extrapolate from evidence of worse risk-adjusted health outcomes for the same groups that social predictors should also predict higher spending. Our findings suggest that adding social factors, particularly race and ethnicity, to the HCC model can entrench health disparities instead of reducing them, by lowering population-based payments to more accurately predicted levels of spending.

Health status, functional status, and access to care were consistently worse than predicted by the HCC model for Black, Hispanic, and less-educated beneficiaries. The lower or similar HCC-adjusted spending for these groups is therefore not explained by better or similar health but, rather, is incommensurate with greater health care needs. Our varied results across groups suggest that moving from fee-for-service to population-based payment under current risk adjustment would reallocate resources to better meet the needs of some groups but not others. HCC-adjusted population payments would increase per beneficiary provider payments for Black and Hispanic beneficiaries by $376–$1,264 (approximately 4–14 percent) above current fee-for-service spending. Although greater increases may be necessary to fully correct underuse and other quality deficits, these are sizable redistributions that require only continued movement away from fee-for-service toward population-based payments. In the case of ACO models, this requires moving away from benchmarks that incorporate historical spending, which reflect underspending for Black and Hispanic beneficiaries, toward a system of risk-adjusted regional rates.[14,15] It is arguably fortuitous that omission of race and ethnicity from the HCC model results in meaningful implicit reallocations insofar as data on race and ethnicity are imperfect; progress need not await better data. Moreover, if more explicit adjustments were

EXHIBIT 4

Differences in access to care, health status, and functional status after standard risk adjustment among fee-for-service Medicare Consumer Assessment of Healthcare Providers and Systems (CAHPS) respondents, 2012–17

	HCC-adjusted difference in access or health measure (percentage points)							
Characteristics	Problem accessing care (mean, 16.2%)	95% CI	Fair or poor general health (mean, 29.3%)	95% CI	Fair or poor mental health (mean, 15.1%)	95% CI	Difficulty with an ADL (mean, 37.5%)	95% CI
RTI race and ethnicity								
Non-Hispanic White	Ref		Ref		Ref		Ref	
Black	5.9	5.3, 6.5	3.0	2.3, 3.7	2.4	1.7, 3.1	1.7	0.9, 2.5
Hispanic	7.0	6.1, 7.9	7.6	6.5, 8.6	5.7	4.9, 6.6	2.0	1.1, 2.9
Education								
High school diploma or more	Ref		Ref		Ref		Ref	
No high school diploma	4.3	3.8, 4.8	10.4	9.9, 10.9	8.4	7.9, 8.8	4.3	3.8, 4.8
Block group measures								
Percent without a high school diploma	0.3	0.1, 0.4	0.3	0.2, 0.5	0.2	0.0, 0.3	0.3	0.1, 0.5
Percent Black	0.9	0.7, 1.1	0.8	0.6, 1.0	0.4	0.3, 0.6	0.8	0.6, 1.1
Percent Hispanic	0.8	0.6, 1.0	1.5	1.2, 1.7	0.9	0.7, 1.1	1.3	1.0, 1.5

SOURCE Authors' analysis of fee-for-service Medicare CAHPS survey data, Medicare claims and enrollment data, and American Community Survey data. **NOTES** All results are derived from analysis of the CAHPS sample. Estimates for each block group measure have been rescaled to reflect a 1-standard-deviation change in the measure. The Research Triangle Institute (RTI) race and ethnicity designation method is described in the exhibit 1 notes. Activities of daily living (ADLs) include bathing, dressing, eating, getting in or out of chairs, walking, and using the toilet. HCC is Hierarchical Condition Categories.

needed, they could face legal challenges, as described by Tim Jost.[16]

In contrast, risk-adjusted population-based payments using the current model would result in minimal reallocations toward beneficiaries with less than a high school diploma. Thus, additional payment reallocations would be needed to better resource efforts to address education-related disparities, which were larger than racial and ethnic disparities in health and functional status. One approach would be to use constrained regression to increase the payment weights on HCCs that are more prevalent among beneficiaries with less education so that payment would exceed current spending for them by a desired amount.[17] An advantage of this approach is that it would not require data for the full population; data on education for a sample of the population (for example, CAHPS) would suffice. Another approach would be to implement post-estimation adjustments (after estimation of the risk-adjustment model) to redistribute payment toward the group of interest. This approach would require data on the full population, which are not currently available at the individual level for education (or many other socioeconomic variables). Accordingly, the ACO Realizing Equity, Access, and Community Health (REACH) model implemented this approach, using community-level variables to increase ACO benchmarks above HCC-predicted spending for communities with greater needs.[18]

When based on analogous sociodemographic predictors at the community level, estimates were generally similar in direction but smaller in magnitude than those based on individual-level predictors. We thus can conclude that moving toward population-based payment under current risk adjustment would not result in substantial within-county payment redistributions between communities with different racial, ethnic, or educational composition. Thus, as for individual-level education, additional payment adjustments would be needed to reallocate resources toward communities in need.

Although payment adjustments at the community level may be considered poorly targeted when the intention is to benefit a subgroup of residents, they may nevertheless be important complements to individual-level adjustments. The latter are critical to establish incentives for MA plans or ACOs to compete for underserved patients. For plans or ACOs to act effectively on those incentives, however, they must include providers serving those patients' communities. Payment reallocations at the area level may have greater influence on market entry and network inclusion decisions made by plans and ACOs.

Furthermore, plans and ACOs may face higher costs relative to fee-for-service spending in historically marginalized communities, including higher transaction costs incurred when contracting with a more fragmented set of providers and the higher costs of developing the necessary in-

Payment adjustments at the community level may be important complements to individual-level adjustments.

formation systems and care management infrastructure to achieve efficiencies under a risk contract. Providers in low-income communities are also less likely to have favorable payer mixes and reserves for financing de novo ACO formation.[19] Thus, community-level incentives may be necessary to encourage formation or entry that might otherwise require distortionary incentives at the individual level (for an illustrative example, see the Supplementary Discussion in the appendix).[13]

Spending for Black, Hispanic, and less-educated beneficiaries remained lower after we removed HCC adjustments from the model (leaving age, sex, enrollment segment, and county as predictors). This finding has two implications. First, it is consistent with less intensive coding of diagnoses for these groups, suggesting that a manipulable risk-adjustment system can undermine the ability of population-based payment to direct resources to populations in need. Second, the finding indicates that spending for Black, Hispanic, and less-educated beneficiaries is lower even without adjustment for clinical conditions. This suggests that substituting non-manipulable social predictors for manipulable diagnoses in risk-adjustment models would, alone, not be sound strategy. Doing so would mitigate coding incentives but also would reintroduce incentives to select favorable clinical risks and eliminate the equity-advancing reallocations implicitly achieved by the current risk-adjustment system. It is critical that efforts to address coding incentives in a population-based payment system be attentive to such trade-offs.

Recommendations For Policy And Research

By departing from predictive accuracy as the singular goal of risk adjustment, a population-based payment system that set payments above current levels of fee-for-service spending for groups with greater deficits in health care access or quality (and set payments below current spending for others) would create incentives for providers or plans to attract those groups and help address resource disparities that contribute to health care disparities. In theory, a combination of intrinsic motivation and competitive pressures should then lead providers or plans to pass the additional resources through to the intended groups in ways that improve their care. On the basis of conceptual consideration and our empirical findings, we anticipate that a combination of additional individual-level and area-level payment adjustments will be needed. Our findings also suggest that the continued expansion of population-based payment models would promote racial and ethnic equity even under the current risk-adjustment system. Ideally, more comprehensive data would be available to support additional individual-level adjustments, but other techniques could be used until such data become available.[17]

The extent to which intended pass-throughs occur is an important topic for research. Studies of Medicare Advantage suggest that competition does indeed promote enhanced offerings to attract enrollees.[20,21] It is unclear, however, whether plans respond to higher payments for historically marginalized groups by offering enhancements for those groups specifically. Our results for fee-for-service spending adjusted for supplemental coverage suggest that MA plans receive payments for Black and Hispanic enrollees that are favorable relative to expected expenditures in the absence of such enhancements. Plans should therefore have strong incentives to compete for them with differential benefits (for example, special supplemental benefits, additional outreach, language-concordant customer service and case management, and broader network breadth for prevalent conditions).

Our fee-for-service estimates are more directly applicable to ACOs than to MA plans. Because provider organizations serve specific communities and exert more direct control over quality of care than do MA plans, ACOs may be positioned to pass through additional payments to specific patient groups in a more targeted fashion. However, ACOs are statutorily limited in the enhancements they can offer, and competition for patients among ACOs may be weaker than that among MA plans—and thus less effective in driving pass-throughs—because patients face higher costs when switching providers than when switching plans. Thus, equity-promoting ACO benchmarks may need to be coupled with mechanisms for passing through tangible benefits directly. For example, Medicare could apply a share of its cut of an ACO's gross savings to reduce Part B and D premiums for the ACO's aligned patients. Patients of ACOs disproportionately serving Black and Hispanic patients, for example, would receive greater premium reductions, all else equal, because of savings induced by setting benchmarks 4–14 percent above spending for Black and Hispanic beneficiaries (as we estimated).

As other indicators of historical disadvantage are considered, coding practices evolve, and risk adjustment is refined (for example, to limit coding incentives), we recommend expanding and repeating our analytic exercise, as the results and implications for payment policy may change. Finally, because the optimal distribution of payment cannot be determined from a predictive exercise, the process must be iterative. As such, it will be important to monitor disparities to understand the impact of initial reallocations and inform subsequent adjustments of population-based payments. ■

This work was supported by grants from the Commonwealth Fund (Grant No. 20223746) and National Institute on Aging, National Institutes of Health (Grant No. P01 AG032952). J. Michael McWilliams reports serving as a senior adviser to the Center for Medicare and Medicaid Innovation, a consultant to RTI International, a consultant to Blue Cross Blue Shield of North Carolina, and an associate editor for *JAMA Internal Medicine*. The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of any of these organizations, the Commonwealth Fund, or the National Institutes of Health.

NOTES

1 Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation. Report to Congress: social risk factors and performance in Medicare's Value-Based Purchasing Program [Internet]. Washington (DC): HHS; 2020 Mar [cited 2022 Nov 17]. Available from: https://aspe.hhs.gov/sites/default/files/migrated_legacy_files/195191/Second-IMPACT-SES-Report-to-Congress.pdf

2 Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447–53.

3 Dieleman JL, Chen C, Crosby SW, Liu A, McCracken D, Pollock IA, et al. US health care spending by race and ethnicity, 2002–2016. JAMA. 2021;326(7):649–59.

4 Wallace J, Lollo A, Duchowny KA, Lavallee M, Ndumele CD. Disparities in health care spending and utilization among Black and White Medicaid enrollees. JAMA Health Forum. 2022;3(6):e221398.

5 Geruso M, McGuire TG. Tradeoffs in the design of health plan payment systems: fit, power, and balance. J Health Econ. 2016;47:1–19.

6 Bergquist SL, Layton TJ, McGuire TG, Rose S. Data transformations to improve the performance of health plan payment methods. J Health Econ. 2019;66:195–207.

7 Better Medicare Alliance. Medicare Advantage risk adjustment: current issues, evolution, and policy recommendations [Internet]. Washington (DC): BMA; 2018 Feb [cited 2022 Nov 17]. Available from: https://bettermedicarealliance.org/wp-content/uploads/2020/03/BMA_RiskAdjustment_WhitePaper_2018_02_27_v2a.pdf

8 Irvin JA, Kondrich AA, Ko M, Rajpurkar P, Haghgoo B, Landon BE, et al. Incorporating machine learning and social determinants of health indicators into prospective risk adjustment for health plan payments. BMC Public Health. 2020;20(1):608.

9 Ash AS, Mick EO, Ellis RP, Kiefe CI, Allison JJ, Clark MA. Social determinants of health in managed care payment formulas. JAMA Intern Med. 2017;177(10):1424–30.

10 Waddill K. Humana calls for social determinants of health in risk adjustment. Health Payer Intelligence [serial on the Internet]. 2020 Apr 23 [cited 2022 Nov 17]. Available from: https://healthpayerintelligence.com/news/humana-calls-for-social-determinants-of-health-in-risk-adjustment

11 Research Data Assistance Center. Research Triangle Institute (RTI) race code [Internet]. Minneapolis (MN): ResDAC; [cited 2022 Nov 17]. Available from: https://resdac.org/cms-data/variables/research-triangle-institute-rti-race-code

12 Jarrín OF, Nyandege AN, Grafova IB, Dong X, Lin H. Validity of race and ethnicity codes in Medicare administrative data compared with gold-standard self-reported race collected during routine home health care visits. Med Care. 2020;58(1):e1–8.

13 To access the appendix, click on the Details tab of the article online.

14 McWilliams JM, Chen A, Chernew ME. From vision to design in advancing Medicare payment reform: a blueprint for population-based payments [Internet]. Washington (DC): USC-Brookings Schaeffer Initiative for Health Policy; 2021 Oct [cited 2022 Nov 17]. Available from: https://www.brookings.edu/wp-content/uploads/2021/10/From-Vision-to-Design-in-Advancing-Medicare-Payment-Reform-1.pdf

15 Chernew M, McGuire T, McWilliams JM. Refining the ACO program: issues and options [Internet]. Boston (MA): Harvard Medical School, Department of Health Care Policy, Healthcare Markets and Regulation Lab; 2014 Nov 7 [cited 2022 Nov 17]. Available from: https://hmrlab.hcp.med.harvard.edu/files/hmrtest2/files/2014_refining_the_aco_program_issues_and_options.pdf

16 Jost TS. Considering race and ethnicity in Covid risk assessments—legal concerns and possible solutions. N Engl J Med. 2022;387(6):481–3.

17 van Kleef RC, McGuire TG, van Vliet RCJA, van de Ven WPPM. Improving risk equalization with constrained regression. Eur J Health Econ. 2017;18(9):1137–56.

18 Centers for Medicare and Medicaid Services. ACO Realizing Equity, Access, and Community Health (REACH) model: request for applications [Internet]. Baltimore (MD): CMS; 2022 Feb 24 [cited 2022 Nov 17]. Available from: https://innovation.cms.gov/media/document/aco-reach-rfa

19 McCullough JM, Coult N, Genau M, Raikhelkar A, Love K, Riley W, et al. Safety net representation in federal payment and care delivery reform initiatives. Am J Accountable Care. 2019;7(1).

20 Song Z, Landrum MB, Chernew ME. Competitive bidding in Medicare Advantage: effect of benchmark changes on plan bids. J Health Econ. 2013;32(6):1301–12.

21 Jacobs PD, Kronick R. The effects of coding intensity in Medicare Advantage on plan benefits and finances. Health Serv Res. 2021;56(2):178–87.